Transit ridership and poverty in urbanized areas: a cross-sectional analysis

Kevin Cox; Curtis Crees; Raghav Kaul

Abstract

Despite theory claiming that transit provides mobility to the lowest-income members of society, empirical evidence for poverty as a determinant for transit ridership is lacking. Most studies of transit ridership encounter two hurdles. Either they simply try to estimate a demand function, relying on individual perceptions of transit utility, without a response variable that accurately reflects demand, or they miss several important explanatory variables for demand itself. This research proposes a novel and statistically robust multivariate regression model to estimate ridership as the intersection of supply and demand, rather than explaining solely a demand curve. It also adds several explanatory variables to existing models of ridership estimation.

I. Introduction

The most basic justification for public transit is providing an acceptable level of mobility for all persons in society (Sanchez 1998). Due to this narrative, in many contexts, public transit transforms from infrastructure to social service. Transit is thereby meant to provide mobility to the lowest income groups, especially in areas stricken with poverty. Given that poverty in urbanized areas throughout the United States is a prevalent reality (Heilman, 2014), one might expect a high level of transit patronage in those urbanized areas.

However, there is a substantial discrepancy between the narrative of transit as social service and its actual function, as the United States is anomalous among western nations in its preference for private vehicle transportation. More broadly underpinning this theory is that perceived utility for public transit is a reflection of the economic laws of consumer demand, but some argue that aggregate ridership is more than the aggregate demand of the entire population. Many scholars have attempted to explain trends in national transit ridership through models of regional geography, economy, population characteristics, and prices of substitutes (primarily private vehicles) (Taylor and Fink 2003, Taylor et al. 2009).

The two questions the literature guides an observer to answer are 1) does transit patronage and investment mean more mobility for the most disadvantaged, access to employment, and therefore lower rates of poverty? 2) If not, what about an urbanized area explains the fluctuation in ridership? Our original analysis attempted to examine the effects of transit ridership on poverty rates, controlling for other variables. Literature and theory pointed to a different causal relationship: demand for ridership was determined by a population's socioeconomic characteristics, not the other way around. A novel model was developed to explain transit ridership among different urbanized areas, the natural unit of analysis for a study on public transit. This model is important to guide future policy decisions on transit investment and evaluate the veracity of conventional narratives on poverty transportation.

II. Lit Review

Investment in public transit is usually made with the implicit assumption that it greatly benefits those in poverty without access to their own transportation, and increases the general welfare by greasing the wheels of labor and capital.(Heilmann 2014) However, whereas European countries exhibit widespread utilization of public transit (Pucher 1988), in the United States, public transit ridership is rare outside the "largest metropolitan areas," (Heilmann 2014), even in the lowest income groups. While these private vehicle substitutes for transit are quite widespread, much attention has been dedicated to the problem of how public transit can provide accommodate increased demand in the lowest income brackets, especially in urbanized areas where fuel taxes, parking, and limitations of the freeway system can hamper the utility of private vehicles (Giuliano 2005).

The literature documents a few reasons for this discrepancy between the supposed poverty-reducing benefits of transit and its virtually universally low patronage. The spatial mismatch hypothesis argues that low-income residents of urban areas are locked out from employment opportunities in the suburbs due to lack of transit options. According to Sanchez (1998), the educational background of workers in urban areas doesn't necessarily match the level of employment offered in those area due to urban sprawl and the de-concentration of jobs from urban centers. This seems to indicate that, controlling for other factors, transit ridership and investment decrease the incidence of poverty, and that "poor workers are more likely to ride public transit than are higher income workers," (Waller 2005). The phenomenon of low-wage and unemployed workers riding public transit is referred to as "poverty transportation."

Despite Sanchez's analysis, the spatial mismatch hypothesis remains complicated to test. While Sanchez argues that ridership in public transit might increase labor force participation and therefore reduce poverty, poverty is supposedly the factor that drives ridership in the first place. Conventional research has experienced difficulty explaining this phenomenon. Furthermore, despite the attention being paid to public transit and urban mobility, "there is very little evidence of the degree to which one affects the other," (Sanchez 1998). This was the basis for the first component of our analysis.

The lack of evidence that "poverty transportation" is widespread in the US leads to the question: who does use public transit? What really drives transit patronage? Despite substantial maintenance and operational expenses in transit, ridership trends have remained stable or declined. Taylor et. al (2009) attempt to answer this question, studying demand for public transportation with a focus on finding the impetus for subsidization. Synthesizing a large amount of previous research, Taylor concludes that previous research "shares little in terms of data, method, or findings," and uses "small sample sizes," two problems this paper attempts to remedy.

The research estimates a demand function for ridership using transit fare, vehicle uptime, service attributes, passenger characteristics, prices of substitutes characteristics of the urbanized area and region, they explain that the "causality arrow between transit service supply and consumption points in both directions," meaning perceived utility of and therefore the demand for transit "varies significantly from person to person and from trip to trip (even for the same person.)" (2009)

Kohn (1999), however, argues that the largest determinant of transit ridership is access to and usage of private vehicles (primarily automobiles). Among important incentives are implicit subsidies to automobiles, population growth, and distance to work. Kohn develops a statistically robust model that explained ridership in terms of trips unlinked trips with vehicle revenue hours (uptime) and average fare.

Both Kohn and Taylor argue that service area population is also among the most significant determinants of ridership. Taylor specifically argues that urbanized area population has an “enormous influence on” transit service supply, and therefore, ridership. Kohn evaluates population by using the two qualitative binary variables: one for cities with populations in excess of one million, one for cities with a service area population lower than 100,000. Kohn later drops the binary variables to yield a statistically robust model with high level of explanatory power.

Kohn’s model forgoes the demand and explains ridership per capita as opposed to estimating demand for transit, the distinction being that demand refers to perceived utility by an individual user and ridership reflects actual consumption - the equilibrium quantity of the supply-demand relationship. Therefore, one can reconcile Kohn's model with the model proposed in Taylor et. al were there a relationship between transit supply and ridership - the primary relationship this paper examines.

To conclude, the discrepancy between poverty transportation and universally low ridership of public transit opens up questions of whether transit patronage can predict the economic misfortunes of entire urbanized areas. Our paper examines these relationships, and then seeks to answer the question that if no such relationship exists, what does determine transit ridership? The important contribution in our analysis is incorporating previously unused variables, as well as establishing the distinction between the demand function of public transit (which is related to the perceived utility by the individual) and the ridership of public transit (which refers to the equilibrium intersection of supply and demand of transit), and finally in estimating the latter of the two.

III. Data

Data Summary Table

Variable	Mean	Standard Deviation	Min	Max
Commute	21.899	4.652	13.3	42.0
Mismatch	21.528	13.405	0.5	63.2
ridershipPC	15.65713	21.412	0.0	229.9
investmentPC	70.903	72.816	1.6	683.8
Pop	594081	1490658	51428	1.84e07
Povrate	17.611	5.598	5.2	37.2
popdensity	2161.345	875.454	811.1	6999.3
Laborforce	73.693	4.851	56.2	84.7

Our initial analysis, predicted a positive relationship between urban poverty and demand for transit (transit ridership), based on the conclusions of Sanchez and Heilman. Any simple regression of poverty would experience substantial omitted variable bias; therefore our multiple regression equation is listed as follows:

$$(1)\ P = \beta_0 + \beta_1 ridershipPC + \beta_2 investmentPC + \beta_3 commute + \beta_4 hsgrad + u$$

Where *P* is the poverty rate of an urbanized area, the unit of our analysis; *ridershipPC* refers to the total unlinked passenger trips as defined in the National Transit Database *per capita*; *investmentPC* refers to total expenditures for FY 2014 as defined in the National Transit Database *per capita*, and commute refers to the mean time to take a trip one-way to employment in minutes. *hschoolgrad* was defined as the percentage of the population who had graduated from high school, and was used as a control. Total ridership and investment figures were found by aggregating individual ridership figures, and per capita figures were found by dividing the aggregates by the population figure listed in the ACS.

Ridership was chosen to represent the utilization of the public transit options in each urbanized area, while investment was chosen as a proxy to represent the robustness of public transit options in the area. The assumption was that high transit utilization indicated higher mobility, especially among the lower income groups, improving access to employment and reducing poverty. The predicted sign on ridership was negative. Each unlinked trip referred to one single ride on bus, train, or light rail, with transfers counting as two or more unlinked trips.

Investment per capita is included, represented by *investmentPC*. The expectation is a negative sign on the variable, consistent with the popular narrative that transit lifts people out of poverty due to access to employment opportunities. Finally, *commute* was chosen to reflect the poverty-increasing effects of slow or limited mobility options, with the expected coefficient being negative.

While Kohn uses vehicle revenue hours to broadly represent "transit availability," revenue hours and miles don't accurately reflect the level of public expenditures on transit services. The use of total investment divided by population as *investmentPC* in our model improves the policy relevance of our analysis for local policymakers making the investment and allocation decisions. Additionally, with Kohn's vehicle revenue hours variable, there was a substantial risk of double-counting from the NTD database as the redundancies in revenue hours are not documented clearly.

Data on poverty levels and commute was retrieved from the American Community Survey 2013 1-year data by urbanized area to most closely reflect the NTD's values from FY 2014. The NTD listed transit data for N = 347 unique urbanized areas while the ACS listed 420 values. Only urbanized areas that were in both the NTD and ACS were used in our analysis. This *N* solves the problem of small sample size mentioned in Taylor et. al (2009).

Indeed, previous research as well as our initial analysis does not point to public transit as a poverty reduction strategy. Instead, some literature tends to explain ridership trends in terms of poverty levels. According to Guiliano, mobility in urbanized areas is a function of resources, meaning that lower-income households who cannot afford private vehicles are likely to be higher patrons, and therefore we hypothesize that higher poverty rates would explain high transit utilization.

The next iterations of the analysis examine the reverse of the relationship from our initial analysis. While a more accurate picture of the response of poverty to transit ridership trends would be painted by examining the time series of one urbanized area, the technique is beyond the scope of this analysis. Our second iteration used a simple regression to estimate the effect of poverty on ridership. Given the relatively low robustness and significance of ridership on poverty from the first OLS regression, little was expected from the newer, reversed OLS regression, defined below:

(2) $R = \beta_0 + \beta_1 povrate + u$

As expected given the number of explanatory variables k =1, our model possesses little explanatory power. Models with a higher level of explanatory power and meaning would have to identify other major contributors to ridership. After examining the works of Kohn and Taylor a more complete model was developed, listed below:

(3) R = β_0+ β_1*povrate* + β_2*commute*+ β_3*mismatch*+ β_4*investmentPC*+ β_5*pop*+ β_6*popdensity*+ β_7*laborforce*+ *u*

Model (3) meets all the Gauss-Markov assumptions for unbiased OLS estimation, fixing issues with earlier models. The inclusion of k = 7 explanatory variables reduces omitted variable bias present in the Model (1)'s overly broad response variable and the Model (2)'s simple regression. No variables are linearly related of any other variables, satisfying the first Gauss-Markov assumption. Our data from the American Community Survey and National Transit Database is the standard dataset for cross-sectional studies, and urbanized areas are assumed to be independently distributed, meeting the assumption of random sampling. Additionally, none of the variables demonstrate substantial multicollinearity.

Correlation table, Model (3)

	povrate	*commute*	*mismatch*	*investmentPC*	*pop*	*popdensity*
povrate	1.000					
commute	-0.3162	1.000				
mismatch	-0.2589	0.5047	1.000			
investmentPC	-0.06	0.2563	0.0417	1.000		
pop	-0.0701	0.3785	0.1233	0.6145	1.000	
popdensity	-0.0345	0.2569	-0.137	0.5092	0.4717	1.000
laborforce	-0.5747	0.0862	0.1468	0.1137	0.1101	0.0399

The sample size of N = 347 guarantees significant sample variation in the explanatory variable, and the reduction of omitted variable bias in Model (3) moves $E(u|x_1...x_k)$ as close to zero as possible with the data. Finally, while large urban metropolises (New York, Los Angeles, Philadelphia) constitute a few notable outliers, the variance of the error u remains roughly constant across all values of $x_1...x_k$. Therefore our model satisfies the Gauss Markov Assumptions for OLS regression.

In this regression, *commute, povrate, R = ridershipPC,* and *investmentPC* refers to the same variables from Model (1). The main additions in Model (3) are *mismatch*, *pop*, *popdensity,* and *laborforce*. Also, *mismatch* refers to the percentage of people in an urbanized area having to commute to a different county for work, as reported in the 2013 1-year ACS figures. The spatial mismatch hypothesis described by Sanchez was the basis for the inclusion of this variable, indicating a high incidence of mismatch between residence and location of available employment should be among the most significant determinants of transit patronage.

The *pop* variable refers to the total population of the urbanized area, included by both Taylor, et. al and Kohn. Population figures are also from the ACS, and are listed by urbanized area. Taylor's argument that population density is among the important demographic determinants of transit demand is the justification for its' inclusion in Model (3), where *popdensity* refers to the residents per square mile in an urbanized area. Finally, labor force participation for adults between the ages of 18 and 65 is included in the variable *laborforce*, based on Sanchez and Walller's argument that populations with larger labor force participation might have enough financial resources to afford themselves private vehicular transportation.

III. Results

a. Model (1)

In Model (1), each explanatory variable was found to be statistically significant at the 1% level, demonstrating a strong set of choices in explanatory variables. The coefficients were also consistent with expectations derived from theory, except for the case of *commute*. Increased transit ridership is correlated with lower poverty rates, meaning high public transit ridership and mobility might be related to the availability of employment opportunities and reduced poverty. Larger investments establish more robust transit options, improving mobility. High school graduation rates were merely a control variable in this regression, though plentiful research indicates other factors require control in any serious examination of poverty as well.

The counterintuitive coefficient for *commute* might be explained by the fact that long commutes indicate overutilization of existing transportation systems, and therefore might indirectly signal higher labor force participation and lower poverty rates.

However, the overall R^2 of Model (1) was low. One explanation is substantial omitted variable bias contained in the error term u, something of an expectation when dealing with models of poverty. There also exists strong multicollinearity between investment and ridership variables, increasing the bias of OLS estimators.

Model (1) Estimation Results

Independent Variables	Model (1) $R^2 = .2936$
ridershipPC	0.1550*** (-6.19)
investmentPC	- 0.0347*** (- 4.64)
commute	- 0.4161*** (-7.24)
hschoolgrad	-.3902*** (-7.72)
intercept	61.0608*** (-12.86)

b. Model (2)

Model (2) attempted to answer the question indirectly posed by the literature: does poverty determine public transit patronage? The expectation was a positive coefficient, alongside the conclusion that impoverished areas are more likely to rely on public transportation for mobility due to a lack of personal vehicles. However, while *povrate* was found to be a significant predictor of ridership *per capita* at a 10% level with the coefficient in the expected direction, the overall model had virtually no explanatory power, boasting an R^2 not significantly different from zero. Such an R^2 is expected with a simple regression. This was the impetus for Model (3).

c. Model (3)

In Model (3), *commute, mismatch, investmentPC, pop* (population)*, popdensity,* and *laborforce* were added to the OLS regression model. All variables were found to be significant at the 5% level with the exception of *mismatch* and *popdensity*, a confusing finding given that spatial mismatch of employment and population density were the two variables the literature found theoretically significant but was unable to test. Possible explanations for the low significance of *mismatch* were that county differences did not correspond to modifications in commute routes or transit availability, and that simply the wrong variable or type of analysis was chosen. The majority of the novelty in our analysis came from understanding the impact of spatial mismatch and population density within urbanized areas on patronage of public transit, however, these variables were found to not be significantly different from zero.

The coefficient on *commute* was the same as predicted in our model, also revealing a counter-intuitive relationship. In urbanized areas with higher mean commute times, ridership of public transit is actually lower. A potential explanation for this phenomenon is that transit is actually a more time-efficient option for commuters, yielding lower average commute times for patrons and therefore driving down the mean commute. Furthermore, labor force participation was mentioned frequently in the literature as a proximal cause of poverty; it was seen as an indication that the population of an urbanized area was taking advantage of the mobility options and had more access to employment opportunities. Instead, our estimation shows labor force participation is negatively correlated with ridership per capita.

The remaining explanatory variables were all significant at a 1% level. Per capita investment in transit was among the most significant predictors of per capita ridership. This result is consistent with the analysis in Kohn in Taylor because aggregate investment stands in for vehicle revenue hours, uptime, trips, service availability service quality, and other "soft" factors that determined overall how robust a transit system was. Investment per capita reflects the expenditures that guarantee all of those factors, and therefore is a suitable proxy.

Total population was also among the strongest predictors of an increase in per capita ridership, consistent with Kohn and Taylor. A possible explanation is that crowded urbanized areas tend to be associated with worse traffic and expectations of a long commute drive people to public transit. Another explanation is that large urbanized areas are more cosmopolitan in nature, and it is the preference of the area residents to use public transit. Alternatively, it might be the transit ridership that allows cities with burgeoning populations to grow, although a time series analysis would be more effective in discerning that relationship.

Poverty rate *povrate* was also found a significant predictor of per capita ridership, reflecting a substantial improvement in significance from the simple regression, though the coefficient only improves by ~ 0.2. This relationship also is consistent with the one predicted by Sanchez, Guiliano, and Waller. The biggest loss in significance is with the regression intercept, which is no longer significant at the 5% level. Model (3) met all the Gauss-Markov assumptions specified previously, as well as not having any large issues with multicollinearity.

d. Model (4)

One final regression was conducted, eliminating the statistically insignificant spatial mismatch variable and the population density variables to yield a less powerful but overall more meaningful model. The goal with Model (4) was to retain significant explanatory variables without sacrificing the strong R^2, and explanatory power, of Model (3). Our OLS estimation for Model (4) yielded results for coefficients similar to those found in Model (3). However, we did find that the explanatory power of *povrate* dropped while the explanatory power of *commute* increased. This indicates that *mismatch* and *laborforce* obscure the relationship of *commute* on *ridershipPC* while artificially strengthening that of the *povrate*. The model itself is described below, excluded from the data analysis section as it was a model virtually interchangeable with Model (3) depending on preference for level of fit. It is followed by OLS estimation results for Models (2), (3), and (4).

(4) $R = \beta_0 + \beta_1 povrate + \beta_2 commute + \beta_3 investmentPC + \beta_4 pop + u$

Model (2), (3), (4) Estimation Results

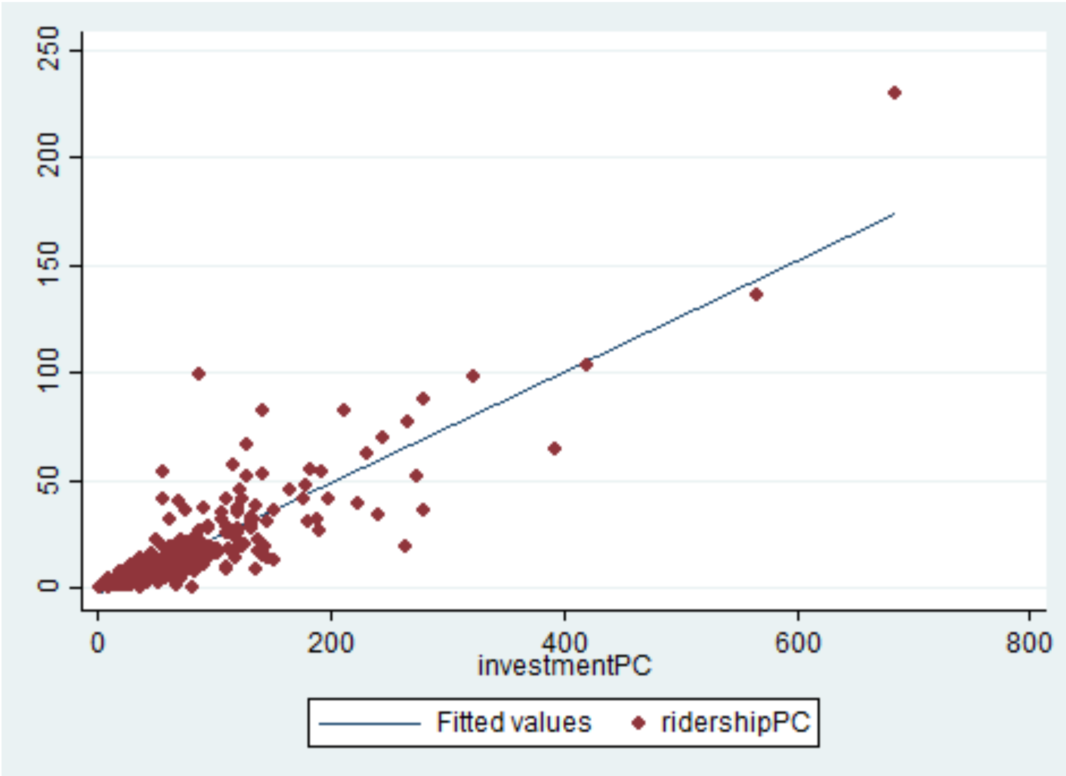
Independent Variables	**Model (2)** $R^2 = 0.0078$, $N = 347$	**Model (3)** $R^2 = 0.8166$, $N = 347$	**Model (4)** $R^2 = 0.8135$, $N = 347$
commute		- 0.4115*** (-2.87)	- 0.2544** (-2.08)
mismatch		0.0739 (-1.61)	
investmentPC		0.228*** (24.88)	0.229*** (26.37)
pop		2.76e-06*** (6.08)	2.80e-06*** (6.28)
povrate	0.3383* (1.65)	0.4027*** (5.33)	0.5029*** (3.47)
popdensity		0.0009 (1.21)	
laborforce		- 0.2156* (1.68)	
constant	9.6999** (2.56)	12.198 (- 1.05)	-5.572 (-1.56)

The F statistics of the model indicate a very significant amount of joint significance amongst the variables in all the models. This allows us to rather soundly reject the notion that the variables in our model have no effect on salary.

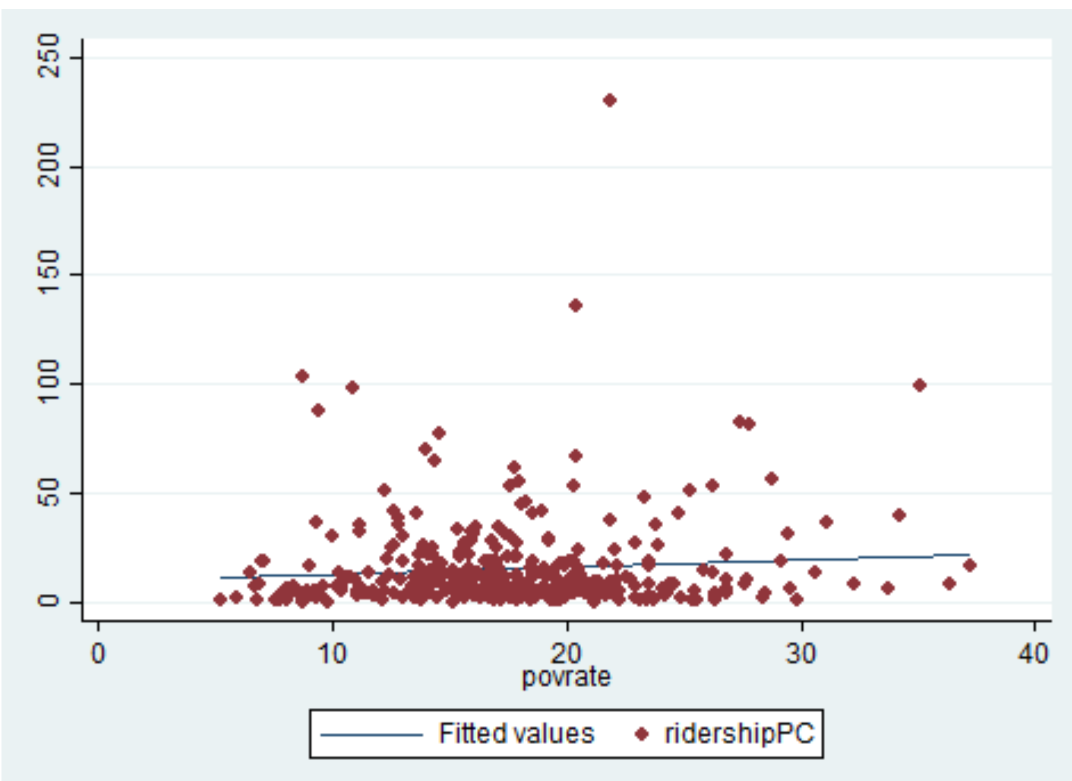
Robustness Tests

	Model (1)	**Model (2)**	**Model (3)**	**Model (4)**
F-test Value	35.54	2.72	215.61	372.02

Ridership per Capita and Investment per Capita



Ridership per Capita and Poverty Rates



IV. Conclusion

There is a strong, convincing correlation between investment in public transportation and its usage, which answers the most fundamental policy question we set out to address – public policy aimed at increasing ridership does indeed fulfill its purpose. The far more interesting question of whether or not this increase in ridership causes significant changes in the poverty rate, however, remains ambiguous, and was perhaps unanswerable within the scope of this paper. Model (1) lacked an especially significant amount of explanatory power in explaining poverty rates in terms of ridership, which is to be expected using such a limited model, and as the enormously varied range of scholarship on poverty rates would suggest, explaining the complex issue of American poverty in terms of her equally complicated transit system was beyond the scope of our model or perhaps indeed any single model.

Estimating transit ridership is not solely a question of estimating demand for public transit. The primary distinction in the literature that guided the statistical testing was one between the estimation of the demand (perceived utility) function and one that estimates ridership. A more heterogenous estimation includes supply-side concerns, primarily investment. We find that the inclusion of other explanatory variables such as poverty rate and investment per capita improves the robustness of the model.

References

Giuliano, Genevieve. 2005. "Low income, public transit, and mobility." *Transportation Research Record: Journal of the Transportation Research Board* no. 1927 (1):63-70.

Heilmann, Kilian. 2014. "Can Public Transport Investment Relieve Spatial Mismatch?".

Pucher, John. 1988. "Urban Travel Behavior as the Outcome of Public Policy: The Example of Modal-Split in Western Europe and North America." *Journal of the American Planning Association* no. 54 (4):509-520. doi: 10.1080/01944368808976677.

Sanchez, Thomas W. 1998. "The Connection Between Public Transit and Employment." *Portland. Portland: Portland State University Center for Urban Studies*.

Taylor, Brian D., and Camille N. Y. Fink. 2003. "The factors influencing transit ridership: A review and analysis of the ridership literature." *University of California Transportation Center*.

Taylor, Brian D., Douglas Miller, Hiroyuki Iseki, and Camille Fink. 2009. "Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas." *Transportation Research Part A: Policy and Practice* no. 43 (1):60-77. doi: 10.1016/j.tra.2008.06.007.

Waller, Margy. 2005. High Cost Or High Opportunity Cost?: Transportation and Family Economic Success. Brookings Institution, Center on Children and Families.

Appendix: Stata Output

```
. regress povrate ridershipPC investmentPC commute hschoolgrad

      Source |       SS       df       MS              Number of obs =     347
-------------+------------------------------           F(  4,   342) =   35.54
       Model |  3184.09427     4  796.023568           Prob > F      =  0.0000
    Residual |   7660.4919   342  22.3990991           R-squared     =  0.2936
-------------+------------------------------           Adj R-squared =  0.2853
       Total |  10844.5862   346  31.3427346           Root MSE      =  4.7328

------------------------------------------------------------------------------
     povrate |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 ridershipPC |   .1550323   .0250498     6.19   0.000     .1057612    .2043034
investmentPC |  -.0347418   .0074947    -4.64   0.000    -.0494834   -.0200002
     commute |  -.4161142   .0574779    -7.24   0.000    -.5291689   -.3030596
 hschoolgrad |  -.3902428   .0505454    -7.72   0.000    -.4896618   -.2908238
       _cons |   61.06082   4.748413    12.86   0.000     51.72105    70.40059
------------------------------------------------------------------------------

. regress ridershipPC povrate

      Source |       SS       df       MS              Number of obs =     347
-------------+------------------------------           F(  1,   345) =    2.72
       Model |  1240.85342     1  1240.85342           Prob > F      =  0.1000
    Residual |  157387.974   345  456.197026           R-squared     =  0.0078
-------------+------------------------------           Adj R-squared =  0.0049
       Total |  158628.827   346  458.464819           Root MSE      =  21.359

------------------------------------------------------------------------------
 ridershipPC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     povrate |   .3382624   .2051019     1.65   0.100    -.0651452      .74167
       _cons |   9.699912   3.789716     2.56   0.011     2.246056    17.15377
------------------------------------------------------------------------------

. regress ridershipPC commute mismatch investmentPC pop povrate popdensity labor
> force

      Source |       SS       df       MS              Number of obs =     347
-------------+------------------------------           F(  7,   339) =  215.61
       Model |  129533.776     7  18504.8252           Prob > F      =  0.0000
    Residual |  29095.0512   339  85.8261098           R-squared     =  0.8166
-------------+------------------------------           Adj R-squared =  0.8128
       Total |  158628.827   346  458.464819           Root MSE      =  9.2642
```

```
------------------------------------------------------------------------------
 ridershipPC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     commute |  -.4115337   .1433912    -2.87   0.004    -.6935822   -.1294852
    mismatch |    .073871   .0459433     1.61   0.109    -.0164988    .1642408
investmentPC |   .2279795   .0091613    24.88   0.000     .2099593    .2459997
         pop |   2.76e-06   4.54e-07     6.08   0.000     1.87e-06    3.65e-06
     povrate |   .4027372   .1162021     3.47   0.001     .1741693    .6313051
  popdensity |   .0008676   .0007177     1.21   0.228    -.0005442    .0022793
  laborforce |   -.215639   .1283784    -1.68   0.094    -.4681575    .0368796
       _cons |   12.19815   11.66984     1.05   0.297    -10.75627    35.15257
------------------------------------------------------------------------------

. regress ridershipPC commute investmentPC pop povrate

      Source |       SS       df       MS              Number of obs =     347
-------------+------------------------------           F(  4,   342) =  373.02
       Model |  129049.282     4  32262.3206           Prob > F      =  0.0000
    Residual |  29579.5452   342  86.4898983           R-squared     =  0.8135
-------------+------------------------------           Adj R-squared =  0.8113
       Total |  158628.827   346  458.464819           Root MSE      =     9.3

------------------------------------------------------------------------------
 ridershipPC |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     commute |  -.2544628   .1223318    -2.08   0.038    -.4950803   -.0138453
investmentPC |    .229677   .0087086    26.37   0.000     .2125478    .2468062
         pop |   2.80e-06   4.45e-07     6.28   0.000     1.92e-06    3.67e-06
     povrate |   .5029008   .0942936     5.33   0.000     .3174323    .6883693
       _cons |  -5.572411   3.565124    -1.56   0.119    -12.58474    1.439919
------------------------------------------------------------------------------
```